

Task 1

InsightSpark wants to access our data to perform some analysis across Australian supermarkets. Before this can happen, I need you to perform some analysis on the provided CSV data set (a sample of the database) to answer the following questions:

- *Across locations, how many apples were purchased in cash?*
- *How much total cash was spent on these apples?*
- *Across all payment methods, how much money was spent at the Bakershire store location by non-member customers?*

Task 2

- Anonymize the provided data set, ensuring the protection of personal details.
- Utilize anonymization techniques such as removing unnecessary columns, masking identifying information, and categorizing sensitive figures.
- Submit the anonymized data as a CSV file.

Task 3

- Explore @CommBank Twitter account.
- Examine the Twitter API's data dictionary.
- Create a proposal outlining insights from @CommBank Twitter data for InsightSpark.

Task 4

- Propose a structured database for storing tweets from the CommBank Twitter account, replies, quote retweets, and direct mentions.
- Define the tables, including the fields, primary keys, and relationships.
- Create a comprehensive database structure in a text submission.